

# Performance of the Sleep-Mode Mechanism of the New IEEE 802.16m Proposal for Correlated Downlink Traffic

Koen De Turck, Stijn De Vuyst, Dieter Fiems, Sabine Wittevrongel and Herwig Bruneel

Ghent University, Department of Telecommunications and Information Processing,  
St. Pietersnieuwstraat 41, B-9000 Ghent, Belgium.

**Abstract.** There is a considerable interest nowadays in making wireless telecommunication more energy-efficient. The sleep-mode mechanism in WiMAX (IEEE 802.16e) is one of such energy saving measures. Recently, Samsung proposed some modifications on the sleep-mode mechanism, scheduled to appear in the forthcoming IEEE 802.16m standard, aimed at minimizing the signaling overhead. In this work, we present a performance analysis of this proposal and clarify the differences with the standard mechanism included in IEEE 802.16e. We also propose some special algorithms aimed at reducing the computational complexity of the analysis.

## 1 Introduction

Great hopes are pinned on the IEEE 802.16e standard (WiMAX) [1], which is poised to be the next big standard in wireless communications. It regulates the communication between mobile stations (MS) and base stations (BS) in a metropolitan area wireless network. Energy-saving mechanisms in wireless communications are currently a hot topic. Short battery life is one of the main impediments to a more widespread use of wireless devices. Hence, understandably, a lot of research is directed at solving or at least mitigating this problem. In the first place, this can be done by making more efficient batteries, but lately there is also a lot of interest in including energy-saving measures in the communication protocols themselves. On that account, it is no wonder that the WiMAX committee has opted to incorporate various energy-saving elements which are commonly referred to as ‘sleep mode’ and ‘idle mode’. Power saving is generally achieved by turning off parts of the MS in a controlled manner when there is neither traffic from the MS (uplink traffic) nor to the MS (downlink traffic). Whereas a MS in sleep mode is still registered to a BS and still performs hand-off procedures, idle mode operation (which is optional in current WiMAX standards) goes further and allows the MS to be completely switched off and unregistered with any BS, while still receiving broadcast traffic.

In this paper, we mainly consider the sleep-mode mechanism, in which the MS turns itself off for predetermined periods of time which are negotiated with

the BS. For the new upcoming IEEE 802.16m standard [2], two companies have proposed a new sleep-mode mechanism, Samsung [3] and Nokia Siemens Networks [4]. The difference between the customary WiMAX sleep-mode mechanism and the new proposals is that the former must be renegotiated every time the traffic volume is low, while in case of the latter mechanisms, traffic is organized in a structure of cycles, and when the system goes into sleep mode, it does so for the rest of the cycle. This requires very few signaling messages. As the Nokia Siemens proposal is a special case of the one championed by Samsung, we will mainly consider the latter.

Sleep-mode operation has received quite some attention lately from the performance modeling community. In [5], the average energy consumption of the MS is obtained in case of downlink traffic only, as well as an approximate expression for the mean packet delay. The energy consumption of the MS in case of both downlink and uplink traffic is considered in [6]. Both [5] and [6] model the incoming (and outgoing) traffic as a Poisson process. An accurate assessment of the delay experienced at the BS buffer however, requires a queueing model. For IEEE 802.16e, in [7] the BS buffer is modeled as a continuous-time finite-capacity queue with a Poisson arrival process and deterministic service times. A semi-Markov chain analysis leads to expressions for the mean packet delay and the mean energy consumption by the MS. The analysis in [8] is based on an  $M/G/1/K$  queueing model with multiple vacations and exhaustive service, where the vacations represent the sleep intervals. Similar work can also be found in [9], where the length of a vacation is assumed to depend on the previous vacation length. In [10], the sleep-mode operation in Cellular Digital Packet Data (CDPD) services is evaluated. The difference with IEEE 802.16e is that the subsequent sleep intervals do not increase in length. The system can thus be modeled as a queueing system with multiple vacations and exceptional first vacation. The loss probability in both [8] and [10] is obtained as well. A simulation study of CDPD sleep-mode performance is found in [11]. An alternative to the exponential increase of the sleep interval lengths is evaluated by simulation in [12]. Finally, in our previous work on the sleep-mode mechanism [14] [13], we considered a general D-BMAP arrival process, and we found that traffic correlation, which was hitherto neglected in almost every study, has an important influence on the sleep-mode performance.

Along with every other paper published on this topic, we assume that the base station has a dedicated buffer for each mobile station that is connected to it, as well as a dedicated amount of bandwidth. It is thus sufficient to consider only a single mobile user. In our opinion, the advantage of modeling a multi-user scenario does not outweigh the significant increase in complexity. The only advantage we see is that one might model to some degree the typical variations in bandwidth that are observed when the number of mobile active users is variable, but this is already a challenging and ongoing research topic without any form of sleep-mode mechanism.

The structure of this paper is as follows. In Section 2, we describe Samsung's proposal and design a mathematical model for it. The analysis is spelled out

in Section 3, where we reduce the problem to a matrix-analytical problem. We develop a computationally efficient variant in Section 4, specifically tailored to this model. We employ our analysis for a WiMAX scenario in Section 5 together with some numerical results. Finally, we draw some conclusions in Section 6.

## 2 Model

Let us first sketch the sleep-mode mechanism that is included in the current IEEE 802.16e protocol. Whenever there is no data to be transmitted between the mobile station (MS) and the base station (BS), the MS can negotiate to go into sleep mode, that is, it will switch off its radio interface for some time, so that the BS cannot transmit packets during this time. There are three types of sleep mode foreseen in the protocol. The simplest is the so-called class III sleep mode, in which a single sleep interval is negotiated. For class II sleep mode, the MS checks at the end of the first sleep interval if the BS has already packets to be transmitted to the said MS, if not, then automatically a new sleep interval of the same length is started. The last, and perhaps the most well-known is the class I sleep-mode mechanism, which is very similar to class II, except that the subsequent sleep intervals do not stay the same but double each time in length until a certain maximum is reached.

In principle, sleep-mode parameters must be renegotiated every time when there is no data traffic between BS and MS (i.e. at the end of each busy period, in queueing parlance). This results in a certain signaling overhead, and it is this shortcoming that the new proposal attempts to address. To this end, data transmission is organized in cycles, consisting of a number of frames. When at a certain point during the cycle, the traffic between BS and MS is interrupted, then a so-called *closedown* interval is started, and if during this closedown interval no new traffic appears, then the mobile station goes into sleep mode until the end of the current cycle. If packets arrive during the closedown interval, then they are transmitted without delay. The advantage of this scheme is that once this sleep mode is in order, it does not have to be renegotiated.

We model the buffer in the base station (from now on denoted as simply ‘the buffer’) as a discrete-time queue with infinite capacity and a first-come-first-served (FCFS) server discipline. This buffer is fed by the traffic in the downlink direction (i.e. from the base station to the mobile station). In this work, as in several other papers on sleep-mode operation, we do not take into account the uplink traffic. We adhere to queueing terminology and denote by a *slot*, what is in the IEEE community usually referred as a frame. The durations of a cycle, of a listening interval at the beginning of a cycle, and of a closedown interval are denoted by  $N$ ,  $L$ , and  $C$  respectively. In our analysis, we will further assume that  $0 \leq L \leq C < N$ , which is conform to most practical situations. Packets arrive at the buffer according to a discrete-time batch Markovian arrival process (D-BMAP). That is, the arrival process is determined by a finite-state Markov chain (with  $M$  states), which is usually referred to as the background Markov chain. The arrival process is defined by the values  $a(k, j|i)$ ;  $k \geq 0$ ;  $i, j \in \{1, \dots, M\}$ ,

denoting the probability that if the background state is  $i$  during a slot, there are  $k$  arrivals during this slot and the background state during the next slot is  $j$ . We put these probabilities in a matrix generating function  $\mathbf{A}(z)$  with dimension  $M \times M$ , which is defined as follows:

$$\mathbf{A}(z) \doteq \sum_{k=0}^{\infty} \mathbf{A}_k z^k, \text{ where } [\mathbf{A}_k]_{ij} \doteq a(k, j|i). \quad (1)$$

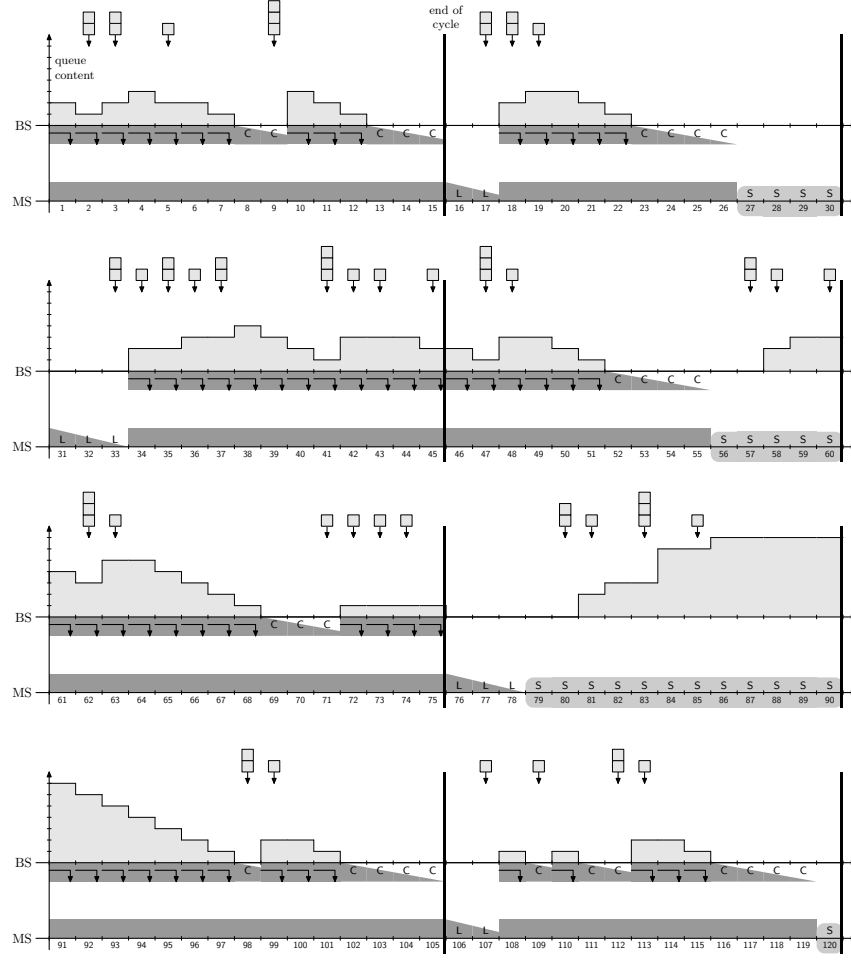
Delayed access is assumed, i.e. arriving packets do not enter the buffer until the end of their arrival slot. Packets are transmitted one by one according to a FCFS discipline. The model at hand differs from a normal single-server queueing model when the buffer gets empty, as in that case a closedown interval of duration  $C$  is started. When during the closedown interval new packets arrive at the BS side, then the closedown interval is abandoned and the packets are served during the next slot. Otherwise, the system goes into sleep mode for the rest of the cycle, and newly arriving packets accumulate in the buffer, waiting to be transmitted during the next cycle. If the closedown interval gets interrupted at the beginning of the next cycle, it gets abandoned as well. At the beginning of each cycle there is a listening interval of  $L$  slots, which behaves identical to a closedown interval. An example of how the system might evolve under the model as described above is depicted in Fig. 1.

The difference between the proposal of Samsung[3] and the one of Nokia Siemens Networks[4] is that the latter does not have a closedown mechanism, and hence corresponds to the special case  $C = 0$ . We conclude this section with some terminology: slots during which a transmission takes place are called ‘active’. Furthermore we distinguish listening slots, closedown slots, and sleep slots, with obvious semantics.

### 3 Analysis

We determine the probability generating function  $U(z)$  of the buffer content during an arbitrary slot. The first step in the analysis is the choice of a convenient set of embedded points and the corresponding Markovian state description. After some trial and error, we settled for the following embedded points: either the beginning of the first slot of a cycle, or the end of a departure slot, that is, a slot in which the packet is transmitted. As there is a fairly large number of possible transitions, we will walk through them, one by one. The Markov state space consists of the buffer content  $u$ , the background state  $s$  and the position  $n$  within the cycle, where we denote the beginning of the cycle with 0, and the end of the different slots with  $1, 2, \dots, N - 1$ . In the next paragraphs, we detail all possible transitions from system state  $(u, n, s)$  to  $(u', n', s')$ , along with the corresponding transition probabilities. These steps take a certain amount of care, as ‘off-by-one’ errors are easily committed.

The first case concerns the functioning of the system in active mode : when there are packets available to be sent ( $u > 0$ ), then one of them is sent and the



**Fig. 1.** System evolution in case  $N = 15, C = 4$  and  $L = 3$ .

embedded Markov chain jumps to the next cycle position (that is, we add one modulo  $N$ )

From	To	Probability
$(u, n, s)$	$((u + k - 1, n + 1 \bmod N, s'))$	$[\mathbf{A}_k]_{ss'}$ if $u > 0$

For the second case, we look at the closedown intervals, more specifically at closedown intervals that are interrupted because of a packet arrival. For this to happen, we must start from an empty buffer, then there must be a number of slots  $i$  in which no arrivals occur, followed by a slot during which there is at least one arrival. The next embedded instant is at the end of the slot thereafter.

From	To	Probability
$(0, n, s)$	$(k, n + i + 2, s')$	$[z^{k+1}][\mathbf{A}_0^{i-1}(\mathbf{A}(z) - \mathbf{A}_0)\mathbf{A}(z)]_{ss'}$ if $0 < n < N - i - 2, 0 \leq i < C$

The notation  $[z^k]F(z)$  denotes the coefficient of power series  $F(z)$  in  $z^k$ . Now we turn our attention to situations where the system enters sleep mode. In that case, we again start from an empty buffer ( $u = 0$ ), followed by a period of  $C$  slots in which no arrivals occur. Then we have a sleep interval of  $N - n - C$  slots in which arrivals accumulate. The next embedded instant is at the cycle boundary.

From	To	Probability
$(0, n, s)$	$(k, 0, s')$	$[z^k][\mathbf{A}(0)^C \mathbf{A}(z)^{N-n-C}]_{ss'}$ if $N - C < n < N$

If the cycle boundary is reached during a closedown interval, then it is not continued, but rather a listening interval is started. Hence the embedded process just jumps to the cycle boundary.

From	To	Probability
$(0, n, s)$	$(0, 0, s')$	$[A(0)^{N-n}]_{ss'}$ if $0 < n \leq N - C$

The last two cases concern the listening interval at the beginning of a cycle. As already noted in the previous section, a listening interval functions in exactly the same way as a closedown interval, except that it occurs only at the beginning of a cycle. Due to that reason, we need not consider the interruption of a listening interval by a cycle boundary. This leads to the last two cases:

From	To	Probability
$(0, 0, s)$	$(k, 0, s')$	$[z^k][\mathbf{A}(0)^L \mathbf{A}(z)^{N-L}]_{ss'}$
$(0, 0, s)$	$(k, i + 2, s')$	$[z^{k+1}][\mathbf{A}(0)^i(\mathbf{A}(z) - \mathbf{A}(0))\mathbf{A}(z)]_{ss'}$ if $0 \leq i < L$

Note that the buffer content  $u$  decreases at most by one, i.e. it is skip-free to the left in  $u$ , which means that the Markov chain can be written as an  $M/G/1$  process with level set  $\{u : u \geq 0\}$  and phase set  $\{(n, s) : 0 \leq n < N, 1 \leq s \leq M\}$ . We give a short account of the solution method in the next section, along with some improvements that are tailored to this specific model.

We now derive relations between the buffer content distribution at embedded points and the distributions during closedown, sleep and listening intervals, in order to eventually establish the distribution at an arbitrary time instant. Essentially, we must ‘fill in the gaps’ that can occur between two embedded time instants. Let us first give some definitions. The (row) vector generating function  $\mathbf{U}_n^A(z)$  records the buffer content distribution at embedded time instant  $n$  for each background state. Specifically, the  $i$ th entry of this vector is the partial pgf

$$[\mathbf{U}_n^A(z)]_i = \sum_{k=0}^{\infty} z^k \Pr[u = k, s = i, n, \mathbf{A}]. \quad (2)$$

As active slots are just one type of slots, this means that  $\sum_n \mathbf{U}_n^A(z)$  is a partial generating function, i.e.  $\sum_n \mathbf{U}_n^A(1)\mathbf{1} < 1$ .

By means of the  $M/G/1$ -technique, we will calculate this vector generating function. Likewise, let  $\mathbf{U}_n^{S(j)}(z)$ ,  $\mathbf{U}_n^{C(j)}(z)$  and  $\mathbf{U}_n^{L(j)}(z)$  denote the vector generating functions at the end of the  $j$ th slot of a sleep, closedown and listening interval respectively. From the description of the model, we find the following equations:

$$\mathbf{U}_j^{L(j)}(z) = \mathbf{U}_0^A(0)\mathbf{A}(0)^{j-1}\mathbf{A}(z) \text{ for } 1 \leq j \leq L, \quad (3)$$

$$\mathbf{U}_n^{C(j)}(z) = \mathbf{U}_{n-j}^A(0)\mathbf{A}(0)^{j-1}\mathbf{A}(z) \text{ for } 1 \leq j \leq C, j < n < N, \quad (4)$$

$$\mathbf{U}_n^{S(j)}(z) = \begin{cases} \mathbf{U}_0^A(0)\mathbf{A}(0)^L\mathbf{A}(z)^j, & \text{for } n = j + L, 0 < j < N - L, \\ \mathbf{U}_{n-j-C}^A(0)\mathbf{A}(0)^C\mathbf{A}(z)^j & \text{for } j + C < n < N, j > 0. \end{cases} \quad (5)$$

By taking the summation of the generating functions over all types of slots, we find the following generating function  $U(z)$  for the buffer content.

$$\begin{aligned} U(z) = & \sum_{n=0}^{N-1} \mathbf{U}_n^A(z)\mathbf{1} + \mathbf{U}_0^A(0)(\mathbf{X}^{(L)}(0) + \mathbf{A}(0)^L\mathbf{X}^{(N-L-1)}(z))\mathbf{1} \\ & + \sum_{j=1}^{N-C-1} \mathbf{U}_j^A(0)(\mathbf{X}^{(C)}(0) + \mathbf{A}(0)^C\mathbf{X}^{(N-C-j)}(z))\mathbf{1} \\ & + \sum_{j=N-C}^{N-1} \mathbf{U}_j^A(0)\mathbf{X}^{(N-j)}(0)\mathbf{1}, \end{aligned} \quad (6)$$

where  $\mathbf{1}$  denotes a column vector with all entries equal to 1 and matrix generating function  $\mathbf{X}^{(k)}(z)$  is defined as

$$\mathbf{X}^{(k)}(z) = \sum_{i=1}^k \mathbf{A}(z)^i = \mathbf{A}(z)(\mathbf{I} - \mathbf{A}(z)^k)(\mathbf{I} - \mathbf{A}(z))^{-1}. \quad (7)$$

We apply the normalization condition  $U(1) = 1$ . Apart from the buffer behavior, the other quantity that is important for a proper evaluation of the sleep-mode mechanism is the expected power consumption  $E$  over a slot. Let external parameters  $E_A$ ,  $E_L$ ,  $E_C$ ,  $E_S$  denote the power consumption during an active, listening, closedown and sleeping slot respectively, then we find that

$$\begin{aligned} E = & E_A \sum_n (\mathbf{U}_n^A(1)\mathbf{1} - \mathbf{U}_n^A(0)\mathbf{1}) + \mathbf{U}_0^A(0) \left( E_L \mathbf{X}^{(L)}(0) + E_S \mathbf{A}^{(L)}(0) \mathbf{X}^{(N-L-1)}(1) \right) \mathbf{1} \\ & + \sum_{j=1}^{N-C-1} \mathbf{U}_j^A(0) (E_C \mathbf{X}^{(C)}(0) - E_S \mathbf{A}^{(C)}(0) \mathbf{X}^{(N-C-j)}(1)) \mathbf{1} \\ & + E_C \sum_{j=N-C}^{N-1} \mathbf{U}_j^A(0) \mathbf{X}^{(N-j)}(0) \mathbf{1}, \end{aligned} \quad (8)$$

## 4 Computational method

In the previous section, we showed that, by choosing the set of embedded points carefully, we can reduce the problem to an  $M/G/1$ -type Markov chain, which has a transition matrix with the following block structure:

$$\mathbf{P} = \begin{pmatrix} \mathcal{B}_0 & \mathcal{B}_1 & \mathcal{B}_2 & \cdots \\ \mathcal{A}_0 & \mathcal{A}_1 & \mathcal{A}_2 & \cdots \\ & \mathcal{A}_0 & \mathcal{A}_1 & \mathcal{A}_2 & \cdots \\ & & \mathcal{A}_0 & \mathcal{A}_1 & \mathcal{A}_2 & \cdots \\ & & & \ddots & \ddots & \ddots \end{pmatrix}. \quad (9)$$

The block matrices  $\mathcal{A}_k$  and  $\mathcal{B}_k$  can be found using the tables in the previous section. Numerical solutions of this type of Markov chain revolve since the pioneering work of Neuts [15] around the computation of the fundamental matrix  $\mathbf{G}$ , which is the smallest nonnegative solution of the matrix equation

$$\mathbf{G} = \sum_{k=0}^{\infty} \mathcal{A}_k \mathbf{G}^k \quad (10)$$

The stationary probability vector  $\boldsymbol{\pi} = [\pi_0, \pi_1, \dots]$  that satisfies  $\boldsymbol{\pi} = \boldsymbol{\pi} \mathbf{P}$ , can subsequently be found by the equations

$$\pi_0 = \pi_0 \mathcal{B}_0^* \text{ and } \pi_n = \left( \pi_0 \mathcal{B}_n^* + \sum_{i=1}^{n-1} \pi_i \mathcal{A}_{n-i+1}^* \right) (\mathbf{I} - \mathcal{A}_0^*)^{-1}, \text{ for } n > 0, \quad (11)$$

where

$$\mathcal{A}_{n+1}^* = \sum_{i=n+1}^{\infty} \mathcal{A}_i \mathbf{G}^{i-n-1}, \quad \mathcal{B}_n^* = \sum_{i=n}^{\infty} \mathcal{B}_i \mathbf{G}^{i-n}. \quad (12)$$

These last expressions are known as Ramaswami's formula, after another pioneer of the matrix-analytic method. There exist a lot of algorithms that aid in computing the desired quantities of interest. Even a straightforward translation of formula (10) constitutes a solid and numerically stable computational method, especially when the Markov chain possesses a strong negative drift. There exist more sophisticated methods as well, for an overview see e.g. [16]. For all these methods, the computational complexity increases very fast with respect to the size of the phase set (here equal to  $MN$ ). Indeed, as we discern a lot of matrix-matrix multiplications of size  $MN$  in the above formulas, the complexity will most often be cubic in  $MN$ , that is  $O(M^3 N^3)$ . As the cycle length  $N$  can be substantial, this would seem bad news for the scalability of our approach. However, the structure that is present in the block matrices  $\mathcal{A}_k$ , allow us to formulate variants of the customary algorithms that are significantly more



efficient. Indeed, block matrices  $\mathcal{A}_k$  have the following form:

$$\mathcal{A}_k = \begin{pmatrix} 0 & \mathbf{A}_k & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{A}_k & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & 0 & \mathbf{A}_k \\ \mathbf{A}_k & 0 & \cdots & 0 & 0 \end{pmatrix} = \mathbf{C}_N \otimes \mathbf{A}_k, \quad (13)$$

where  $\mathbf{C}_N$  is the cyclic permutation matrix that maps the vector  $(v_1, v_2, \dots, v_{N-1}, v_N)$  to  $(v_N, v_1, v_2, \dots, v_{N-1})$ . Hence the matrix  $\mathcal{A}_k$  is block circulant, which means that there exists a block row vector  $\mathbf{a}$  such that  $[\mathcal{A}_k]_{ij} = [\mathbf{a}]_{i-j \bmod N}$ . In this case, the block vector  $\mathbf{a}$  is equal to  $[0, \mathbf{A}_k, 0, \dots, 0]$ . Apart from being efficiently storable (only the first block row or block column must be stored), block circulant matrices are closely related to the discrete Fourier transform, and that fact speeds up the basic operations considerably.

Consider two block circulant matrices  $\mathbf{M}_1$  and  $\mathbf{M}_2$  with first block rows equal to resp.  $\mathbf{m}_1$  and  $\mathbf{m}_2$ . The product  $\mathbf{M}_1 \mathbf{M}_2$  is again block circulant [19] with first block row equal to  $\text{IDFT}(\text{DFT}(\mathbf{m}_1) * \text{DFT}(\mathbf{m}_2))$ , where  $\text{DFT}(\cdot)$  and  $\text{IDFT}(\cdot)$  denote the discrete Fourier transform and the inverse discrete Fourier transform respectively, and the operation  $*$  denotes the component-wise product of the blocks: let  $\mathbf{c} = [\mathbf{C}_1, \dots, \mathbf{C}_N]$  and  $\mathbf{d} = [\mathbf{D}_1, \dots, \mathbf{D}_N]$  denote two block vectors of length  $N$ , with  $M \times M$  block sizes, then  $\mathbf{c} * \mathbf{d}$  is equal to

$$\mathbf{c} * \mathbf{d} = [\mathbf{C}_1 \mathbf{D}_1, \dots, \mathbf{C}_N \mathbf{D}_N]. \quad (14)$$

In order to derive an efficient procedure to compute the fundamental matrix  $\mathbf{G}$ , we first note that because the matrices  $\mathcal{A}_k$  are block circulant,  $\mathbf{G}$  is block circulant as well. This can be intuited from the stochastic interpretation of the matrix  $\mathbf{G}$ . Entry  $[\mathbf{G}]_{ij}$  denotes the probability that, starting from level  $n$  in phase  $i$ , the process returns to level  $n - 1$  in finite time, and that  $j$  is the first phase visited in level  $n - 1$ . This probability does not depend on the cycle positions at the beginning and end of such excursions, but merely on the difference (modulo  $N$ ) between the cycle positions.

All this means we can perform the iterative computation of  $\mathbf{G}$  entirely in the DFT domain. Indeed, let  $\hat{\mathbf{g}}$  denote the DFT transform of the first block row of  $\mathbf{G}$ , and let  $\hat{\mathbf{a}}_k$  denote the DFT of  $[0, \mathbf{A}_k, 0, \dots, 0]$ , then we find from relation (10):

$$\hat{\mathbf{g}} = \sum_{k=0}^{\infty} \hat{\mathbf{a}}_k * \hat{\mathbf{g}}^{*k} \quad (15)$$

This leads to an iterative procedure whose computational complexity is much better than that of the customary algorithms, as the  $*$ -operation costs  $NM^3$  floating point operations, a huge contrast with the  $N^3 M^3$  floating point operations of the original iteration. This variant of the standard M/G/1 technique may have applications beyond the ones in this paper.

There are some faint correspondences between the technique developed in this section and the paper by Gail et al. [17] for structured Markov chains with limited displacement, in that in both cases there is a two-tier block structure, and the fundamental matrix has a special structure as well, although in our case it is block circulant and in the latter case it is the power of a block companion matrix. The use of DFT transforms in the matrix-analytic method became popular with Meini's work [18].

## 5 Practical Example

In this section, we show some examples of what is possible with the model developed in the present article. In all examples, we will use an interrupted Poisson arrival process, conform to guidelines of the IEEE community [21] for the 802.11 standard. This is a two-state (on/off) Markov model with Poissonian arrivals with rate  $\lambda$  in the on-state and no arrivals in the off-state. The Markov chain spends a fraction  $\sigma$  of the time in the on-state. The coefficient of correlation between the states in two subsequent slots is  $1 - \frac{1}{K}$ , where  $K$  is a measure for the mean lengths of both on and off periods, given by  $\frac{K}{1-\sigma}$  and  $\frac{K}{\sigma}$  respectively. It can easily be verified that the matrix generating function  $\mathbf{A}(z)$  for this traffic model equals

$$\mathbf{A}(z) = \begin{pmatrix} e^{\lambda(z-1)} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 - \frac{1-\sigma}{K} & \frac{1-\sigma}{K} \\ \frac{\sigma}{K} & 1 - \frac{\sigma}{K} \end{pmatrix}. \quad (16)$$

The case of uncorrelated states from slot to slot clearly corresponds to  $K = 1$ . Larger values of  $K$  imply that on-slots and off-slots occur more clustered together in time and the arrival process is more bursty. The correlation parameter  $K$  is therefore also referred to as the burstiness factor or the burst-length factor of the arrival process [24], [25]. As transmission times have a fixed length of 1 slot, the load  $\rho$  is equal to  $\rho = \lambda\sigma$ .

For the power consumption during the different types of slots we use the values summarized in Table 5. As there are as yet no consumer-end WiMAX devices, we have used the corresponding WiFi data [23] as a guideline. The performance measures we will look at are the mean packet delay and the mean power consumption.

**Table 1.** Power consumption parameters

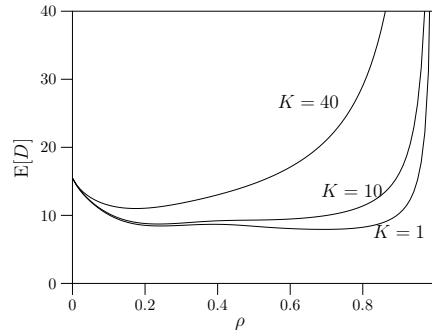
Parameter	Value	Description
$E_S$	0.045 W	Sleep mode power
$E_L$	1.15 W	Listening mode power
$E_C$	1.15 W	Power during closedown
$E_A$	1.65 W	Transmit mode power

In the first pair of plots, Fig. 2 and Fig. 3, we look at the influence of the traffic correlation on the mean packet delay and mean power consumption. We

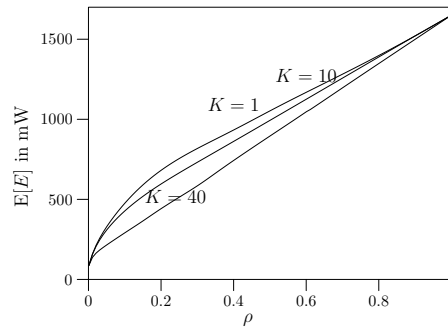
see a confirmation of the well-known fact that traffic correlation deteriorates the queueing performance tremendously. Also note that the system performance gets worse for very light loads. On the other hand, correlation has a small but marked improvement on the average power consumption. Also note that the delay gets larger for very light loads, which is intuitively clear from the fact that in that case the system goes into sleep mode for longer time periods, and hence incoming packets must wait for a longer time.

The next pair of plots, Fig. 4 and Fig. 5 shows the influence of the length  $C$  of the closedown interval on the performance measures, for loads  $\rho = 0.2$  and  $\rho = 0.4$ . As can be expected, larger values of  $C$  signify an increase of the power consumption and a decrease of the mean delay.

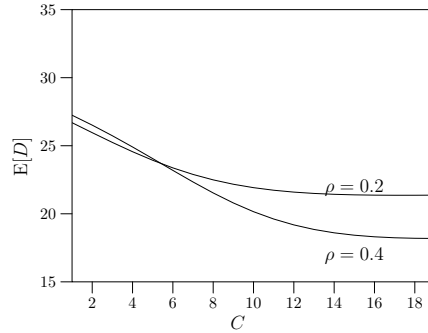
We see the same story, but to a larger degree, in the last two plots, Fig. 6 and Fig. 7, where we show the influence of the cycle length on the performance  $N$ .



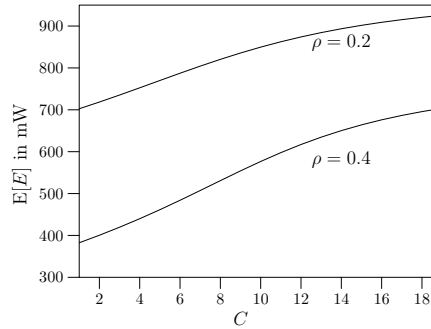
**Fig. 2.** Mean packet delay versus the load for different levels of correlation, where  $N = 32$ ,  $L = 1$ ,  $C = 8$ , and  $\sigma = 0.2$ .



**Fig. 3.** Energy consumption versus the load for different levels of correlation, where  $N = 32$ ,  $L = 1$ ,  $C = 8$ , and  $\sigma = 0.2$ .



**Fig. 4.** Mean packet delay versus the closedown time for different values of the load, where  $N = 64$ ,  $L = 1$ , and  $\sigma = 0.2$ .



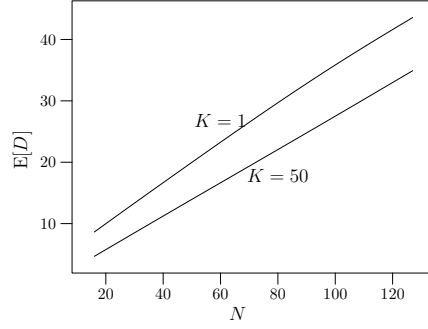
**Fig. 5.** Energy consumption versus the closedown time for different values of the load, where  $N = 64$ ,  $L = 1$ , and  $\sigma = 0.2$ .

## 6 Conclusion

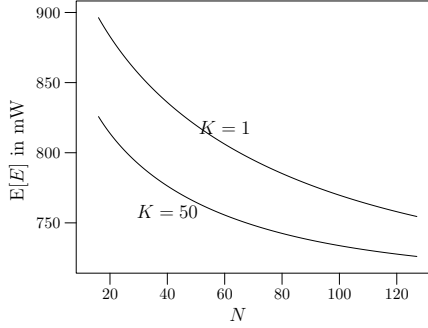
We analyzed Samsung's proposal of a new sleep-mode mechanism to be included in the next version of IEEE 802.16m WiMAX protocol, by means of a queueing model. From the preliminary numerical results, we see that the proposal shows the same qualitative behavior as the customary (802.16e) sleep-mode mechanism, and thus provides the same room for optimization of the parameters  $C$ ,  $L$  and  $N$ . This optimization forms the focus of our planned further work.

## Acknowledgements

The third author is a Postdoctoral Fellow with the Research Foundation–Flanders (FWO–Vlaanderen), Belgium.



**Fig. 6.** Mean packet delay versus the cycle length for different levels of correlation, where  $L = 1, C = 8, \rho = 0.2, \sigma = 0.2$ .



**Fig. 7.** Energy consumption versus the cycle length for different levels of correlation, where  $L = 1, C = 8, \rho = 0.2, \sigma = 0.2$ .

## References

1. IEEE 802.16e-2005, "Part 16: Air interface for fixed and mobile broadband wireless access systems — Amendment 2: physical and medium access control layers for combined fixed and mobile operation in licensed bands — Corrigendum 1," February 2006.
2. The Draft IEEE 802.16m System Description Document, IEEE 802.16m-08/003r4, July 2008.
3. Sleep mode operation for IEEE802.16m, C802.16m-08/721r1, July 2008.
4. Keep-awake mechanism for 802.16m sleep mode, C802.16m-08/718, July 2008.
5. Y. Xiao, "Energy saving mechanism in the IEEE 802.16e wireless MAN," *IEEE Communications Letters*, Vol. 9, No. 7, 2005, pp. 595–597.
6. Y. Zhang and M. Fujise, "Energy management in the IEEE 802.16e MAC," *IEEE Communications Letters*, Vol. 10, No. 4, 2006, pp. 311–313.
7. K. Han and S. Choi, "Performance analysis of sleep mode operation in IEEE 802.16e mobile broadband wireless access systems," *Proceedings of the IEEE 63rd Vehicular Technology Conference, VTC2006-Spring* (Melbourne, 7-10 May 2006), Vol. 3, pp. 1141–1145.

8. Y. Park and G.U. Hwang, "Performance modelling and analysis of the sleep-mode in IEEE 802.16e WMAN," *Proceedings of the IEEE 65th Vehicular Technology Conference*, VTC2007-Spring (Dublin, 22-25 April 2007), pp. 2801–2806.
9. J.-B. Seo, S.-Q. Lee, N.-H. Park, H.-W. Lee, and C.-H. Cho, "Performance analysis of sleep mode operation in IEEE 802.16e," *Proceedings of the 60th Vehicular Technology Conference*, VTC2004-Fall (Los Angeles, 26-29 September 2004), Vol. 2, pp. 1169–1173.
10. S.-J. Kwon, Y.W. Chung, and D.K. Sung, "Queueing model of sleep-mode operation in cellular digital packet data," *IEEE Transactions on Vehicular Technology*, Vol. 52, No. 4, 2003, pp. 1158–1162.
11. Y.-B. Lin and Y.-M. Chuang, "Modeling the sleep mode for cellular digital packet data," *IEEE Communications Letters*, Vol. 3, No. 3, 1999, pp. 63–65.
12. N.-H. Lee and S. Bahk, "MAC sleep mode control considering downlink traffic pattern and mobility," *Proceedings of the IEEE 61st Vehicular Technology Conference*, VTC2005-Spring (Stockholm, 30 May–1 June 2005), Vol. 3, pp. 2076–2080.
13. K. De Turck, S. De Vuyst, D. Fiems, and S. Wittevrongel, "An analytic model of IEEE 802.16e sleep mode operation with correlated traffic," *Proceedings of NEW2AN 2007*, pp. 109–120.
14. K. De Turck, S. De Vuyst, D. Fiems, and S. Wittevrongel, "Performance analysis of the IEEE 802.16e sleep mode for correlated downlink traffic," *Telecommunication Systems*, Vol. 39, No 2, pp. 145–156.
15. M.F. Neuts, "Structured stochastic matrices of  $M/G/1$  type and their applications", volume 5 of *Probability: Pure and Applied*. Marcel Dekker Inc., New York, 1989.
16. D.A. Bini, G. Latouche, B. Meini, "Numerical Methods For Structured Markov Chains", Oxford University Press, 2005.
17. H. R. Gail, S. L. Hantler, B. A. Taylor. "Non-skip-free  $M/G/1$  and  $G/M/1$  type Markov chains" *Advances in Applied Probability*, Vol. 29, No. 3 (Sep. 1997), pp. 733–758.
18. B. Meini. "An improved FFT-based version of Ramaswami's formula". *Stochastic Models*, 13(2):223–238, 1997.
19. C. Van Loan. "Computational frameworks for the fast Fourier transform", volume 10 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.
20. L. Berlemann, C. Hoymann, G.R. Hiertz, and S. Mangold, "Coexistence and interworking of IEEE 802.16 and IEEE 802.11(e)," *IEEE 63rd Vehicular Technology Conference (VTC)*, 1:27–31, 2006.
21. C.R. Baugh, J. Huang, R. Schwartz, and D. Trinkwon, "Traffic model for 802.16 tg3 mac/phy simulations," *Technical report, IEEE 802.16 Broadband Wireless Access Working Group*, 2001.
22. D. Sivchenko, B. Xu, V. Rakocevic and J. Habermann, "Internet traffic performance in IEEE 802.16 networks," *European Wireless*, European Wireless 2006, Athens, Greece, 2-5 April 2006.
23. E.-S. Jung and N.H. Vaidya, "Improving IEEE 802.11 power saving mechanism," *Wireless Networks*, Vol. 14 No. 3 pp. 375–391, 2004.
24. S. Wittevrongel, H. Bruneel, *Discrete-time queues with correlated arrivals and constant service times*, *Computers & Operations Research*, Vol. 26, No. 2, 1999, pp. 93–108.
25. S. De Vuyst, S. Wittevrongel, H. Bruneel, *Mean value and tail distribution of the message delay in statistical multiplexers with correlated train arrivals*, *Performance Evaluation*, Vol. 48, No. 1–4, 2002, pp. 103–129.